Piecewise linear trend detection in mesosphere/lower thermosphere wind time series

R. Q. Liu and Ch. Jacobi

Summary

A piecewise linear model is developed to detect climatic trends and possible structural changes in time series with a priori unknown number and positions of breakpoints. The initial noise is allowed to be interpreted by the first- and second-order autoregressive models. The goodness of fit of candidate models, if the residuals are accepted as normally distributed white noise, is evaluated using the Schwarz Bayesian Information Criterion. The uncertainties of all modeled trend parameters are estimated using the Monte-Carlo method. The model is applied to the mesosphere/lower thermosphere winds obtained at Collm (52°N, 15°E) during 1960-2007. A persistent increase after ~1980 is observed in the annual mean zonal wind based on the primary model while only a weak positive trend arises in the meridional component. Major trend breakpoints are identified around 1968-71 and 1976-79 in both the zonal and meridional winds.

1. Introduction

As with global change near the Earth's surface, there is also of interest to detect long-term trends in the upper atmosphere and attribute them to their primary causes. Recently, a relatively consistent pattern of middle and upper atmosphere temperature trends has been presented, showing cooling in the stratosphere/mesosphere, weak trend around the mesopause, and cooling in the thermosphere (Laštovička et al., 2008). However, when dynamical parameters in the middle and upper atmosphere are considered, a much less clear picture is found. Now available mesosphere/lower thermosphere (MLT) wind time series of more than three decades indicate that wind trends may be interrupted, or change direction (Portnyagin et al., 2006; Jacobi et al., 2009; Merzlyakov et al., 2009).

These changes in trends may be analysed using statistical models. Unlike in some pioneer structural change trend analyses, e.g. on the turnaround and recovery of the total ozone column (Reinsel et al., 2002) or changes of the global surface temperature anomaly (Seidel and Lanzante, 2004), where the possible change dates are specified in advance, the number and times of possible trend breaks in MLT winds are a priori unknown (Tome and Miranda, 2004), i.e., they can only be determined according to some basic mathematical principles that underpin the proposed model. This also increases the complexity and skill needed from a practical algorithm. In addition, an integral trend model should be able to not only detect possible trend breakpoints (BPs) and measure the associated partial trends but also as fully as possible account for the variability of an original time series. So an implicit fundamental assumption for a statistical model is that the ultimate modeled errors need (or can) not be explained any more, e.g., the residuals can be regarded as independent and identically distributed (i.i.d.) random variables with zero mean and common variance (Reinsel et al., 2002; Seidel and Lanzante, 2004).

A statistical model for structural change trend assessment, incorporating the methods proposed by Tome and Miranda (2004) and Seidel and Lanzante (2004), has been developed. It will be applied to analyze the climatic trends and their structural changes in the mid-latitude MLT wind series obtained at Collm (52°N, 15°E) during 1960-2007.

2. A piecewise linear trend model

As a natural extension of the linear regression model, let us consider the following structural change linear regression model with *m* BPs T_1, T_2, … T_m (and thus *m*+1 regimes or segments) applied to a time series of the length *T*:

$$Y_t = \sum_{i=1}^{m+1} I_{\{T_{i-1}+1 \le t \le T_i\}} \left(a_i + b_i t\right) + N_t, \quad (t = 1,\ 2,\ \ldots,\ T), \tag{1}$$

where $T_0 = 0$, $T_{m+1} = T$. I_A denotes an indicator variable equal to one if the event *A* is true (e.g. when $t \in [T_{i-1}+1,\ T_i]$) and zero otherwise (e.g. when $t \notin [T_{i-1}+1,\ T_i]$). A continuity condition at each turning point is imposed as

$$a_i + b_i T_i = a_{i+1} + b_{i+1} T_i, \quad (i = 1,\ 2,\ \ldots,\ m). \tag{2}$$

In Eq. (1), Y_t is the observed dependent variable at time t, a_i and b_i $(i = 1,\ 2,\ \ldots,\ m+1)$ are the corresponding trend regression coefficients (i.e. intercept and slope) for each segment, and N_t is the unexplained noise term often assumed to be autoregressive with time lag of 1 or 2 (AR(1) or AR(2), e.g. Reinsel et al., 2002; Seidel and Lanzante, 2004). That is, $\{N_t\}$ satisfies $N_t = \varphi N_{t-1} + \varepsilon_t$ or $N_t = \varphi_1 N_{t-1} + \varphi_2 N_{t-2} + \varepsilon'_t$, where the errors ε_t (ε'_t) are independent random variables with mean 0 and common variance σ_ε^2 $(\sigma_{\varepsilon'}^2)$ and

$$\varphi = \rho_1,\ \ \varphi_1 = \rho_1(1-\rho_2)/(1-\rho_1^2),\ \ \varphi_2 = (\rho_2 - \rho_1^2)/(1-\rho_1^2)\ ,$$

when assuming $\{N_t\}$ is a stationary random process with standard lag-one and -two autocorrelations ρ_1 and ρ_2.

Note that this is a partial structural change model in the sense that the autoregressive parameters are assumed to be constant across regimes. The BPs T_1, T_2, … T_m are explicitly treated as unknown. Our procedure is first to estimate the unknown piecewise linear trend coefficients together with the times of BPs when *T* observations on Y_t are available. Then the produced noise term will be tentatively interpreted, respectively, by the first- and second-order autoregressive models as well as that one without autoregression (AR(0)) when the N_t themselves can be regarded as independent random errors with zero mean and common variance σ_N^2. Finally, the uncertainties of all modeled trend parameters (including the positions of BPs) are estimated using the Monte-Carlo method.

In general, the number of structural breaks m is also unknown. However, at the beginning, we treat it as known (i.e. apply the procedure with different m) and its determination will be treated later as a problem of model selection. The method of estimation considered is that based on the least-squares principle (Bai and Perron, 1998). For each m-partition $(T_1,\ T_2,\ ...,\ T_m)$, the associated least-squares estimates of trend coefficients are obtained by minimizing the “sum of squared residuals (SSR)” (as in Tome and Miranda (2004), we treat slopes of line segments and intercept of the first segment as the independent regression coefficients and so employ an efficient algorithm proposed therein to create the design matrix):

$$S_T = \sum_{t=1}^{T}\left[Y_t - \sum_{i=1}^{m+1} I_{\{T_{i-1}+1\le t\le T_i\}}\left(a_i + b_i t\right)\right]^2 , \tag{3}$$

and the estimated BPs $\hat{T}_1,\ \hat{T}_2,\ ...,\ \hat{T}_m$ are such that

$$(\hat{T}_1,\ \hat{T}_2,\ ...,\ \hat{T}_m) = \arg\min_{T_1,\ ...,\ T_m} S_T\left(T_1,\ T_2,\ ...,\ T_m\right), \tag{4}$$

where the minimization is taken over all partitions $(T_1,\ T_2,\ ...,\ T_m)$ subject to a set of appropriate constraints on the minimum distance between two consecutive BPs, the minimum length for the first and last segments and the minimum amount of trend change at BPs (Tome and Miranda, 2004; 2005).

In practice, one can start with the case of zero BP (i.e. the simple linear case when Eqs (1) and (3) are still valid but (2) and (4) disappear naturally), up to a maximum of $m\ (\ge 1)$ BPs. For each of the $m+1$ cases the following step is to augment the corresponding regression trend with the first- and second-order autoregressive components. As did in Seidel and Lanzante (2004), we assess the goodness of fit of the residuals (hereafter i.e. the modeled errors) to a one-dimensional (1-D) Gaussian distribution, both with removal of the AR(1) or AR(2) behavior in the noise and directly with the model AR(0), by using the Anderson-Darling (AD, e.g. Romeu, 2003) statistic to test the null hypothesis of normally distributed residuals. We eliminate from further consideration any model for which the null hypothesis is rejected at the 5% significance level (see Table 1A in Stephens, 1974). On the other hand, the mean and the standard lag-one and -two autocorrelations of each accepted normally distributed residual series are calculated to check whether it can be regarded as a realization of a white noise process. Only after this we, in principle, employ the standard form of the Schwarz Bayesian Information Criterion (BIC, Ng and Perron, 2005; Portnyagin et al., 2006):

$$S(q) = T\ln\left[\frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2\right] + q\ln T , \tag{5}$$

where $\hat{Y}_t$ denotes the modeled value (vs. the residual) of the dependent variable at time t and $q = 2m+2$ for AR(0), $q = 2m+3$ for AR(1) and $q = 2m+4$ for AR(2) (Seidel and Lanzante, 2004), to select the primary/best and secondary models as those with the lowest and second-lowest values of BIC, provided that the residuals are accepted as 1-D normally distributed white noises.

Finally, an important and unavoidable issue is the statistical significance of the estimated BPs and partial trends whereas it is still an open discussion (Tome and Miranda, 2005). For each accepted residual series (hereafter, as a 1-D normally distributed white noise), however, it is reasonable to assume that the residuals are i.i.d. and follow a common distribution $N(0,\sigma_N^2)$ for AR(0), $N(0,\sigma_\varepsilon^2)$ for AR(1) or $N(0,\sigma_{\varepsilon'}^2)$ for AR(2). Thus it is convenient, using the Monte-Carlo simulation approach, to estimate the standard deviations of all modeled trend parameters (one can repeatedly generate the corresponding pseudorandom normally distributed residual series (Press et al., 1992), add it to the modeled sequence of the dependent variable and run the first step of the foregoing procedure, and at last compute the sample mean and variance of all the fitted trend parameters).

3. Application to Collm wind data

The model is applied to Collm MLT zonal and meridional prevailing winds during 1960-2007. The data evaluation and first trend analysis results have already been presented in Jacobi et al. (1997) and Jacobi and Kürschner (2006). There have been several changes in measuring strategy, which can potentially lead to inhomogeneity in the time series and thus to possible artifacts in trend analysis. During the first decade of the measurements, data analysis has been performed manually, with smaller measuring density in the early years. In particular, before 1968 data have been only taken during the evening hours, so that these years cannot be regarded as reliable in a trend analysis. The switch from manual to automatic data analysis in 1972 has been accompanied by a very long (several years) parallel analysis, so that artifacts due to this change are improbable. The change from the analysis of single time series to an average over three measuring paths is connected with a smoothing of the time series. Therefore, year-to-year variability before and after 1979 may show an apparent change, which is not of meteorological origin. However, the analysis of long-term trends should not be seriously affected.

Because we are mainly concerned about the climatic trends and their structural changes in the MLT winds and to avoid so-called end effects (Tome and Miranda, 2005), the minimum distance between adjacent BPs and the minimum length for the first and last segments are both set to 5 years in this study. The allowed minimum amount of trend change at BPs is 0.01 ms^{-1}/year. These constraints are optimized for our problem and changing them moderately would not have a significant effect on the modeled results. To accurately estimate the standard deviations of all fitted trend parameters when using the Monte-Carlo method, we always generate 10000 pseudorandom series (actually only ~8300 series are used because ~17% of them are rejected at a 15% significance level through the AD test) to simulate the corresponding normally distributed i.i.d. residuals. Some model parameters and input/output data files are listed in Appendix A.

The model is applied to annual mean winds, which are expected to disclose stable trend results, although one has to keep in mind that annual mean winds in the MLT have limited physical meaning.

Fig. 1 shows annual mean zonal winds with corresponding trends added, based on different pure trend models with 0 BP up to 5 BPs (from bottom to top) but without autoregression. At first, the AD tests (hereafter at the 5% significance level) and related statistic calculations (see Table 1) reveal that only the models with 2 up to 5 BPs can produce acceptable residuals. In other words, both the simple linear assumption and the 1-BP pure trend model (showing results similar to those obtained in Portnyagin et al. (2006) but with a larger variance of the break point time) have to be eliminated from further consideration owing to the non-Gaussian distribution of their residual series and the large lag-one autocorrelations as well. Then from Table 2 we find that the best choice according to BIC is the 2-BP pure trend model (i.e., without autoregressive component). It exhibits 2 major turning points, respectively, in 1971 and 1979, and after that a persistently positive trend (0.22 ms^{-1}/year) arises. Nevertheless, the large wind variability before the late 1970s has not yet been completely removed by annually averaging (refer to Fig. 3 below). This strong variability probably includes some artifacts, and in turn, it will "mislead" the BIC (see the right hand side of Eq. (5)) to select a simpler model having BPs only before ~1980. In this case, as the number of fitting parameters q increases, the second term $q \ln T$ will increase rapidly whereas the first term (proportional to the SSR) decrease slowly (refer to Fig. 3 below), together leading to an increase of the value of BIC. This suggests that in reality the 3- and 4-BP pure trend models should also be considered as acceptable choices (we reject the 5-BP fit, which shows the same times of the last 4 BPs as in the 4-BP model, because of its high value of BIC). This provides 2 additional possible trend breaks, i.e. those in 1991 and 1998/99. These BPs are almost independent/quasi-stable solutions because their respective uncertainty intervals have no evident overlaps with those of other BPs. Or, more properly, they can be regarded as "minor shifts" which are superposed on a persistently increasing background wind after 1979.

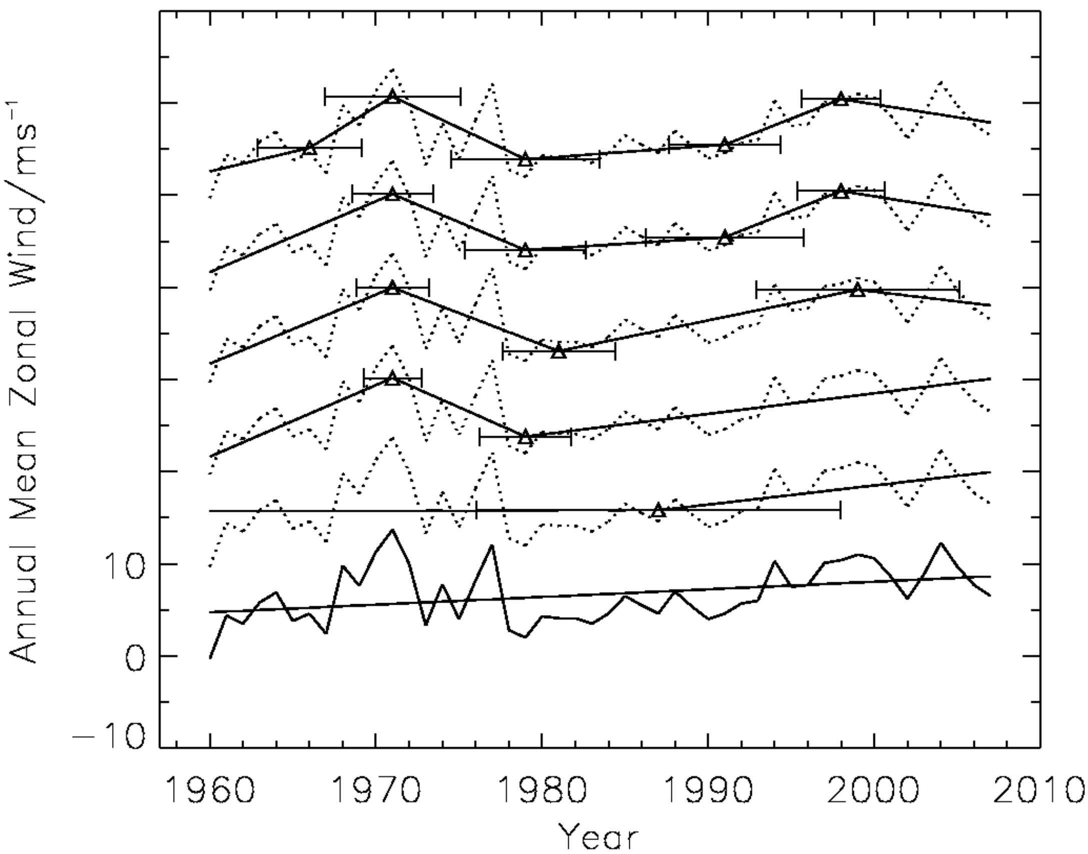

Fig. 1. *Time series of the annual mean zonal wind with corresponding trends added in turn, from bottom to top, based on different pure trend models (with 0 BP up to 5 BPs but without autoregression).*

All the trend break years detected above have, within the limits of their uncertainty, been identified in the winter prevailing zonal wind observed over Obninsk (55°N, 37°E) as well using a sophisticated WZ-method (Merzlyakov et al., 2009). Furthermore, the turnaround at ~1990 has been given particular attention recently in the combined Collm and Obninsk winds, because it may indicate a structural change in trends in dynamics of the whole northern mid-latitude middle atmosphere up to the lower thermosphere (Portnyagin et al., 2006; Jacobi et al., 2009).

Table 1: *Mean (μ), standard lag-one (ρ_1) and -two (ρ_2) autocorrelations, AD-statistic (A^{2*}) and associated significance level (α) of normal distribution testing of each residual series based on different pure trend models (i.e., with m -BP trend but without autoregressive component) applied to the time series of annual mean zonal wind. The number symbols (#) indicate unacceptable residuals at the 5% significance level, but the corresponding statistic values are still listed for comparison.*

AR(0) \ *m*	*0*	*1*	*2*	*3*	*4*	*5*
μ (ms-1)	.00	.00	.00	.00	.00	.00
ρ_1	.42	.35	.12	.06	.02	-.01
ρ_2	.18	.09	-.19	-.26	-.28	-.28
A^{2*}	.90	.83	.38	.28	.50	.61
α	# <.05	# <.05	>.15	>.15	>.15	>.10

Table 2: *Values of BIC based on different pure trend or combination models (i.e., with m -BP trend plus r -order autoregressive component) applied to the time series of annual mean zonal wind. The best model is identified with an asterisk (*). The number symbols indicate cases of unacceptable residuals, but the corresponding BIC values are still listed for comparison.*

r \ *m*	*0*	*1*	*2*	*3*	*4*	*5*
0	# 108.52	# 113.39	* 100.44	104.12	108.90	115.41
1	102.88	110.26	103.56	107.80	112.75	119.28
2	106.35	114.16	105.17	108.02	112.66	# 119.40

As in the analysis for the zonal wind, the annual mean meridional wind and its candidate trends, based on different pure trend models, are presented in Fig. 2. From Tables 3 and 4 we find that only the models with 2 up to 4 BPs produce acceptable residuals and again the best fit, according to BIC, is provided by the 2-BP pure trend model. It exhibits 2 major turning points, respectively, in 1968 and 1975. After that only a weak positive trend (0.06 ms^{-1}/year) arises in the annual mean meridional wind, which is due to the different trends in different seasons (Jacobi and Kürschner, 2006). As is the case with the zonal wind, we suggest that in reality the 3- and 4-BP pure trend models should be considered as reasonable alternatives. This discloses 2 expanded trend breaks, i.e. those in 1981 and 2001 (while adjusts the second major BP from 1975 to 1976), though the first one indicates a large uncertainty interval overlapping with the small one of the major BP in 1976.

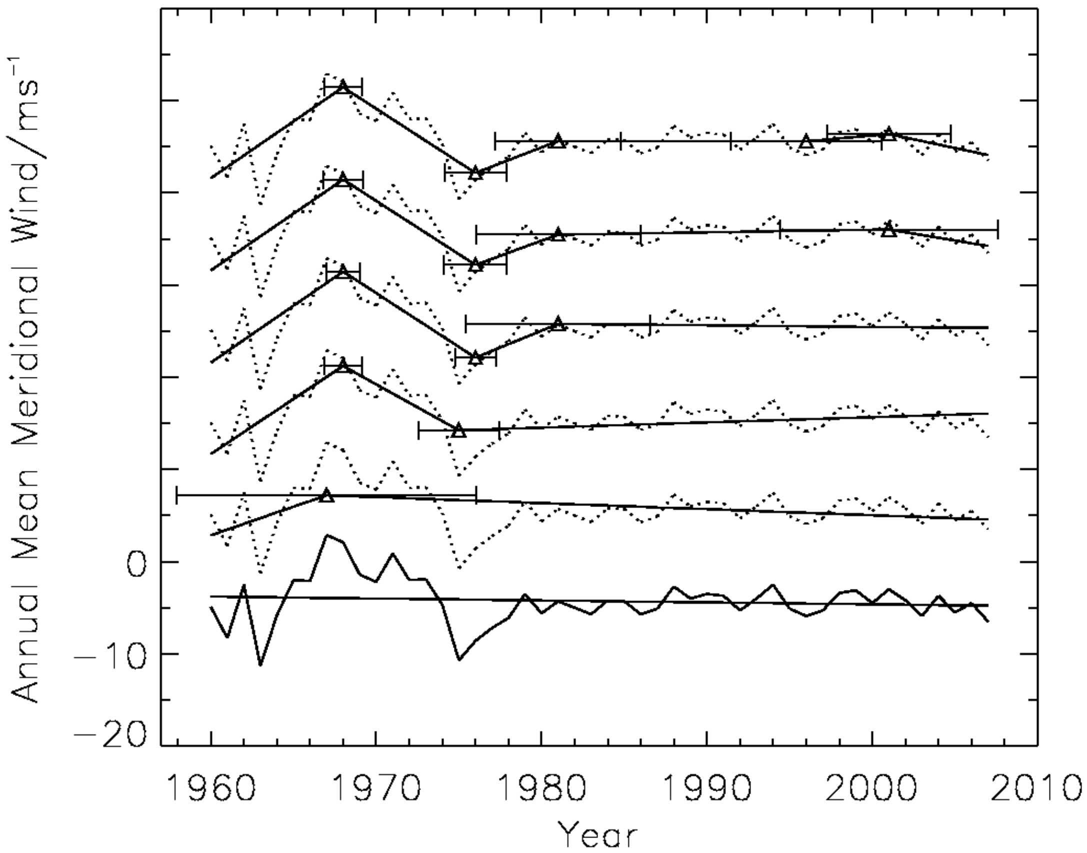

Fig. 2. *Same as Fig. 1 except for the annual mean meridional wind.*

Table 3: *Same as Table 1 except for the annual mean meridional wind.*

AR(0) \ *m*	*0*	*1*	*2*	*3*	*4*	*5*
μ (ms^{-1})	.00	.00	.00	.00	.00	.00
ρ_1	.48	.41	.05	-.07	-.09	-.09
ρ_2	.34	.32	.03	-.04	-.06	-.06
A^{2*}	1.11	.84	.55	.61	.76	.79
α	# <.01	# <.05	>.15	>.10	>.05	# <.05

Table 4: *Same as Table 2 except for the annual mean meridional wind.*

m \ r	*0*	*1*	*2*	*3*	*4*	*5*
0	# 99.51	# 100.98	* 85.74	87.56	93.93	# 101.44
1	# 91.31	# 96.05	89.47	91.17	97.41	104.90
2	# 94.19	# 97.59	93.18	95.08	101.23	108.74

Fig. 3 demonstrates the variations of the estimated SSR (sum of squared residuals) and BIC with different pure trend models applied to the annual mean zonal and meridional winds, respectively. One can see that, compared with the simple linear assumption and the 1-BP case, the 2-BP pure trend model leads to a drastic decrease of the SSR and thus to a sharp drop of the BIC. However, once the BPs assumed in the winds exceed 2, the SSR only decreases slowly so that the BIC turns to increase almost linearly with the increasing number of BPs. Therefore the 2-BP pure trend models obtain the minimum BIC. Nevertheless, as mentioned above, because the large wind variability before the late 1970s probably includes some artifacts and, in turn, contributes to the drastic decrease of the SSR, the 3- and 4-BP pure trend models, which prove to produce acceptable residuals, should in principle be considered as alternative choices.

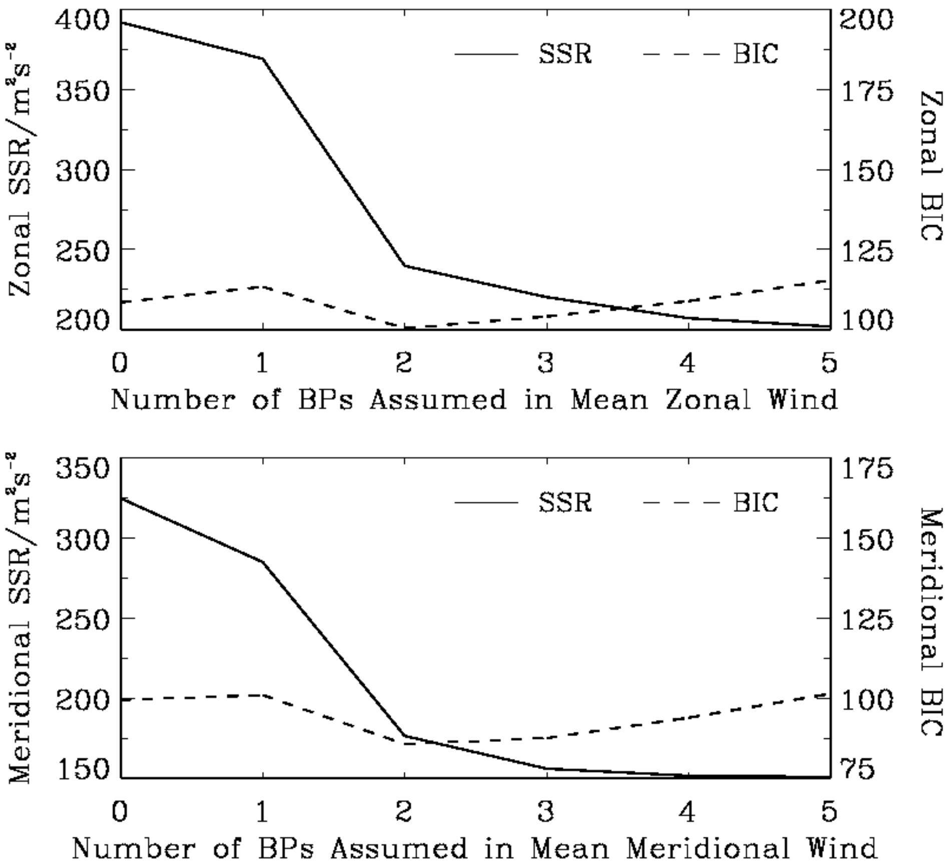

Fig. 3. *Variations of the estimated SSR (solid lines) and BIC (dashed lines) with different pure trend models (with 0 BP up to 5 BPs but without autoregression) applied, respectively, to the annual mean zonal (upper panel) and meridional (lower panel) winds.*

The complete modeling results (modeled series plus residuals) of the annual mean zonal and meridional winds based on the primary and secondary models selected according to BIC (see Tables 2 and 4) are displayed in Fig. 4. One can find that the two fits for the zonal wind are from different (pure trend and combination) structural models since the secondary model has incorporated a first-order autoregressive component, while for the meridional component the two fits are from the same (pure trend) structural models. However, for the zonal wind the reference meaning of the secondary fit is weak because the pure linear trend assumption has proved to be unacceptable (Table 1) and the secondary model has a value of BIC (102.88) much closer to those (103.56 and 104.12) of the third and fourth models than to the BIC (100.44) of the primary model (Table 2). For the meridional component the reference meaning of the secondary fit is strong since the 3-BP pure trend model proves to produce acceptable residuals (Table 3) and the first four models have almost equally spaced BIC values (85.74, 87.56, 89.47 and 91.17, see Table 4).

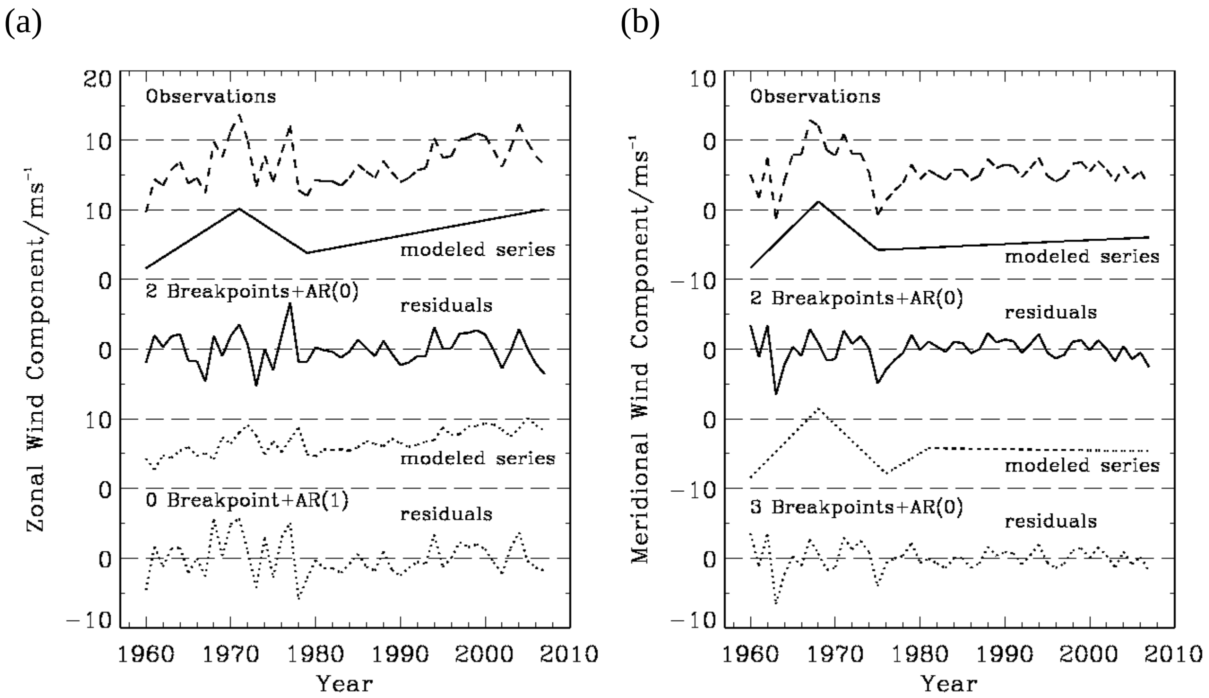

Fig. 4. *Time series of the annual mean (a) zonal and (b) meridional winds (dashed line) and their complete modeling results based on the primary (2 BPs plus AR(0) in both cases: solid lines for the modeled series and the residuals) and secondary (linear plus AR(1) for (a) and 3 BPs plus AR(0) for (b): dotted lines for the modeled series and the residuals) models selected according to BIC.*

4. Discussion

In some cases (e.g. in the seasonal mean winds, not shown here) the initial noises must be further interpreted by an AR(1) or AR(2) even if based on the primary models, suggesting that other unidentified factors or processes may also play a role in determining the evolution of the mesospheric winds. So a multivariate linear regression model (Reinsel et al., 2005) would be a subsequent extension of the basic model in (1) to estimate the effects of other natural factors, possibly including the equatorial Quasi-biennial Oscillation and/or the Southern Oscillation, on the behavior of the MLT winds.

Although attempting to perform the trend analyses as objectively as possible, some subjective decisions unavoidably remain. In particular, when incorporating an AR component in the piecewise linear trend model one always assumes that the initial noise term is a stationary random process. In addition, although the results obtained according to BIC may be statistically robust, there is no unique criterion to select models (Seidel and Lanzante, 2004).

In the case of data showing large local variability in time, it is desirable to consider a heteroskedastic autoregressive component in our model. However, it seems difficult to obtain exact-meaning solutions for all unknown parameters when the form of heteroskedasticity of errors is also unknown, though some statistical-meaning solutions can be modeled, based on the maximum likelihood principle and the use of Gibbs sampler, assuming a WZ-model in which the level, trend and error variance are subject to synchronous structural changes (Wang and Zivot, 2000; Merzlyakov et al., 2009).

We have also only used the annual mean data starting from 1968, i.e. those during 1968-2007 to do the corresponding analyses, and found that the most reasonable fits are from the 1-, 2- and 3-BP pure trend models that produce piecewise linear trends resembling those shown in Figs 1 and 2 based on the 2-, 3- and 4-BP pure trend models but removing the first segments. So there are only 3 trend BPs identified in the zonal and meridional winds with almost same times as those last three shown in Figs 1 and 2, while the first two BPs in 1976 and 1981 in the meridional wind (during 1968-2007) can even be distinguished with independent uncertainty intervals. However, considering that longer data with more samples will generally produce more reliable modeling we only show the results using the data during 1960-2007.

5. Conclusions

A piecewise linear regression model is developed to detect climatic trends and possible structural changes in the time series with a priori unknown number and positions of breakpoints based on the least-squares principle. The initial noise term is allowed to be interpreted by the first- and second-order autoregressive models. In principle, the goodness of fit of candidate models, provided that the modeled residuals are accepted as 1-D normally distributed white noises, is evaluated using the Schwarz Bayesian Information Criterion. The standard deviations of all modeled trend parameters are estimated using the Monte-Carlo method. As an example, this piecewise linear model is applied to the mesosphere/lower thermosphere winds obtained at Collm (52°N, 15°E) during 1960-2007. The main results are as follows:

After ~1980 a persistent increase is observed in the annual mean zonal wind based on the primary model selected according to BIC. During nearly the same period of time, however, only a weak positive trend arises in the annual mean meridional wind due to different trends in different seasons.

Major trend BPs are identified in 1968/71 (maybe physically meaningless because of the possible data artifacts before 1968) and ~1976/79 in the annual mean meridional and zonal winds according to BIC. However, in view of the large wind variability before the late 1970s, the 3- and 4-BP pure trend models, which prove to produce acceptable residuals, should in principle be considered as alternative choices. This provides 4 additional possible minor breaks, i.e. those in 1981, 2001 and in 1991, 1998/99, respectively, in the meridional and zonal winds. In fact, the last three of them are almost independent/quasi-stable solutions, and the first one is even selected by BIC itself as a secondary solution although it indicates a large uncertainty interval overlapping with that small of the major BP in 1976.

Acknowledgments

This research was supported by DFG under grant JA 836/22-1.

References

Bai, J., Perron, P., 1998: Estimating and Testing Linear Models with Multiple Structural changes. *Econometrica, 66*, 47–78.

Jacobi, Ch., R. Schminder, D. Kürschner, J. Bremer, K. M. Greisiger, P. Hoffmann, and W. Singer, 1997: Long-term trends in the mesopause wind field obtained from D1 LF wind measurements at Collm, Germany, *Adv. Space Res.*, *20*, 2085–2088.

Jacobi, Ch., Kürschner, D., 2006: Long-term trends of MLT region winds over Central Europe. *Phys. Chem. Earth*, *31*, 16–21.

Jacobi, Ch., Hoffmann, P., Liu, R. Q., Križan, P., Laštovička, J., Merzlyakov, E. G., Solovjova, T. V., Portnyagin, Yu. I., 2009:, Midlatitude mesopause region winds and waves and comparison with stratospheric variability. *J. Atmos. Sol.-Terr. Phys.*, *71*, 1540–1546.

Laštovička, J., Akmaev, R. A., Beig, G., Bremer, J., Emmert, J. T., Jacobi, C., Jarvis, G. Nedoluha, M. J., Portnyagin, Yu. I., Ulich, T., 2008 : Emerging pattern of global change in the upper atmosphere and ionosphere. *Ann. Geophys.*, *26*, 1255–1268.

Merzlyakov, E. G., Jacobi, Ch., Portnyagin, Yu. I., Solovjova, T. V., 2009: Structural changes in trend parameters of the MLT winds based on wind measurements at Obninsk (55°N, 37°E) and Collm (52°N, 15°E). *J. Atmos. Sol.-Terr. Phys.*, *71*, 1547–1557.

Ng, S., Perron, P., 2005: Practitioners' corner: a note on the selection of time series models. *Oxford Bulletin of Economics and Statistics*, *67*, 115–134.

Portnyagin, Yu. I., Merzlyakov, E. G., Solovjova, T. V., Jacobi, Ch., Kürschner, D., Manson, A., Meek, C., 2006: Long-term trends and year-to-year variability of mid-latitude mesosphere/lower thermosphere winds. *J. Atmos. Sol.-Terr. Phys.*, *68*, 1890–1901.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992: Transformation method: exponential and normal deviates. Numerical recipes in Fortran 77, 2nd ed., Cambridge University Press, Cambridge, pp. 277–280.

Reinsel, G. C., Weatherhead, E. C., Tiao, G. C., Miller, A. J., Nagatani, R. M., Wuebbles, D. J., Flynn, L. E., 2002: On detection of turnaround and recovery in trend for ozone. J. Geophys. Res., 107(D10), 4078, doi:10.1029/2001JD000500.

Reinsel, G. C., Miller, A. J., Weatherhead, E. C., Flynn, L. E., Nagatani, R. M., Tiao, G. C., Wuebbles, D. J., 2005: Trend analysis of total ozone data for turnaround and dynamical contributions, J. Geophys. Res., 110, D16306, doi:10.1029/2004JD004662.

Romeu, J. L., 2003: Anderson-Darling: a goodness of fit test for small samples assumptions. *Reliability Analysis Center*, Selected Topics in Assurance Related Technologies, *10*, 1–6.

Seidel, D. J., Lanzante, J. R., 2004: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *J. Geophys. Res.*, *109*, D14108, doi:10.1029/2003JD004414.

Stephens, M. A., 1974: EDF Statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.*, *69*, 730–737.

Tomé, A. R., Miranda, P. M. A., 2004: Piecewise linear fitting and trend changing points of climate parameters. *Geophys. Res. Lett.*, *31*, L02207, doi:10.1029/2003GL019100.

Tomé, A. R., Miranda, P. M. A., 2005: Continuous partial trends and low-frequency oscillations of time series. *Nonlinear Processes in Geophysics*, *12*, 451–460.
Wang, J., Zivot, E., 2000: A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend and Variance. *Journal of Business and Economic Statistics*, *18*, 374–386.

Appendix A: Model parameters and input/output data files

The model is written in FORTRAN source code. Input and output files and parameters are described in the following.

A1: Input parameters. Only these 8 parameters need to be set appropriately before running the model. Currently this has to be done in the source code. All other parameters have fixed values.

Parameter[*]	*Notation*
ITM	Number of data points of the original time series.
NTS	Actual number of data points extracted from the original time series for the piecewise linear trend analysis.
NEND	Minimum length for the first and last segments set to avoid end effects.
NSPACE	Minimum distance between adjacent BPs.
CSLOPE	Minimum amount of trend change at BPs.
MINCYCLE	Minimum number of Monte Carlo loops set to estimate the standard deviations of all modeled trend parameters.
MULTIPLE	A multiplication factor set to skip Monte Carlo loop numbers between the MINCYCLE and MAXCYCLE.
MAXCYCLE	Maximum number of Monte Carlo loops set to estimate the standard deviations of all modeled trend parameters.

[*] Names of variables in the source code.

A2: Input. There is only one input file. The file contains the data in one column.

Input Data File[*]	*Notation*
AVWINDE.DAT	Original time series.

[*]Currently to be set in the source code (status='old').

A3: Output files.

Output Data File[*]	*Notation*
MBPYEAR.DAT	A 5*6 matrix where the non-zero elements in each column denote the times of BPs.
MSR.DAT	A 3*6 matrix where each column elements denote the minimum SSR obtained when assuming the order of autoregression is equal to the row index.
MQR.DAT	A 3*6 matrix where each column elements denote the minimum BIC obtained when assuming the order of autoregression is equal to the row index.
MCOEFF.DAT	A 7*6 matrix where the non-zero elements of each column denote the (m+2)-element fit vector $\{b_1, b_2, ..., b_{m+1}, a_1\}$, i.e. the slopes of (m+1) segments and the intercept of the first segment.
MEYTS.DAT	A NTS*6 matrix where each column denotes the piecewise linear fit of the original time series without autoregression.
MRYTS.DAT	A NTS*6 matrix where each column denotes the initial noise series obtained without autoregression.
MSTATISTICS.DAT	An 11*6 matrix where each column denotes the statistics of the initial noise series, i.e. sample mean, standard deviation, lag-zero/one/two autocorrelations, standard lag-one (or coefficient for AR(1)) and -two autocorrelations, lag-one and -two coefficients for AR(2), and the AD-statistic and associated significance level at which the null hypothesis of normally distributed residuals is not rejected.
MARMBP.DAT	The best and secondary choices of AR and BPs according to BIC.
MUBPYEAR.DAT	Mean times of BPs obtained via different numbers of loops (MCYCLE) of Monte-Carlo simulations for AR(0) residuals.
STDBPYEAR.DAT	Standard deviations of BPs obtained via different numbers of loops (MCYCLE) of Monte-Carlo simulations for AR(0) residuals.
MUCOEFF.DAT	Means of fit vector $\{b_1, b_2, ..., b_{m+1}, a_1\}$ obtained via Monte-Carlo simulations for AR(0) residuals.
STDCOEFF.DAT	Standard deviations of fit vector $\{b_1, b_2, ..., b_{m+1}, a_1\}$ obtained via Monte-Carlo simulations for AR(0) residuals.
MACSTDBPYEAR.DAT	Similar to STDBPYEAR.DAT but obtained via MAXCYCLE loops of Monte-Carlo simulations.
MACSTDCOEFF.DAT	Similar to STDCOEFF.DAT but obtained via MAXCYCLE loops of Monte-Carlo simulations.

[*]A data matrix of 6 columns in an output file (status='new') assumes the corresponding number of BPs is equal to the column index minus one. Data files including “1” or “2” in the filenames are corresponding to the file without number, but are valid for a combination of the model with an AR(1) or AR(2) component (e.g., MEYTS.DAT, ME1YTS.DAT, ME2YTS.DAT).